Bayesian approach and Naturalness in MSSM analyses for the LHC

M. E. Cabrera [a1], **J. A. Casas** [a2] and **R. Ruiz de Austri** [b3]

[a] *Instituto de Física Teórica, IFT-UAM/CSIC,*
U.A.M., Cantoblanco, 28049 Madrid, Spain

[b] *Instituto de Física Corpuscular, IFIC-UV/CSIC, Valencia, Spain*

Abstract

The start of LHC has motivated an effort to determine the relative probability of the different regions of the MSSM parameter space, taking into account the present, theoretical and experimental, wisdom about the model. Since the present experimental data are not powerful enough to select a small region of the MSSM parameter space, the choice of a judicious prior probability for the parameters becomes most relevant. Previous studies have proposed theoretical priors that incorporate some (conventional) measure of the fine-tuning, to penalize unnatural possibilities. However, we show that such penalization arises from the Bayesian analysis itself (with no ad hoc assumptions), upon the marginalization of the $\mu-$parameter. Furthermore the resulting effective prior contains precisely the Barbieri-Giudice measure, which is very satisfactory. On the other hand we carry on a rigorous treatment of the Yukawa couplings, showing in particular that the usual practice of taking the Yukawas "as required", approximately corresponds to taking logarithmically flat priors in the Yukawa couplings. Finally, we use an efficient set of variables to scan the MSSM parameter space, trading in particular B by $\tan\beta$, giving the effective prior in the new parameters. Beside the numerical results, we give accurate analytic expressions for the effective priors in all cases. Whatever experimental information one may use in the future, it is to be weighted by the Bayesian factors worked out here.

---

[1]maria.cabrera@uam.es
[2]alberto.casas@uam.es
[3]rruiz@ific.uv.es

1 Introduction

The imminent start of LHC has motivated an interesting effort (see refs. [1–8]) to antici-
pate which kind of supersymmetric model is more likely to be there, or, in more precise
words, which region of the parameter space of the minimal supersymmetric standard
model (MSSM) is more probable, taking into account the present (theoretical and ex-
perimental) wisdom about the model. This wisdom includes theoretical constraints (and
perhaps prejudices) and experimental constraints, such as electroweak precision tests.
The idea is to use this information to determine the relative probability of the different
regions of the MSSM parameter space, thus the frequent expression "LHC forecasts".
The appropriate framework to evaluate this probability is the Bayesian approach, which
allows to separate in a neat way the objective and subjective pieces of information.

In the Bayesian analysis one tries to make inferences about the relative probability of
different "states of nature" (corresponding to different values of the parameters defining
the model, say p_i) upon the observation of different data which are determined[4] completely
by p_i.

The probability density of a particular point $\{p_i^0\}$ in the parameter space, given a
certain set of *data*, is the so-called posterior probability density function (pdf), $p(p_i^0|\text{data})$,
which is given by the fundamental Bayesian relation (for a review see ref. [9])

$$p(p_i^0|\text{data}) \;=\; p(\text{data}|p_i^0) \; p(p_i^0) \; \frac{1}{p(\text{data})} \; . \tag{1.1}$$

Here $p(\text{data}|p_i^0)$ is the likelihood (sometimes denoted by $\mathcal{L}$), i.e. the probability density of
measuring the given data for the chosen point in the parameter space. E.g. for observables
measured within a gaussian uncertainty, $\mathcal{L}$ is proportional to $e^{-\frac{1}{2}\chi^2}$, where χ^2 is the
conventional chi-squared. $p(p_i^0)$ is the prior, i.e. the "theoretical" probability density that
we assign a priory to the point in the parameter space. Finally, $p(\text{data})$ is a normalization
factor which plays no role unless one wishes to compare different classes of models, so for
the moment it can be dropped from the previous formula.

One can say that in eq. (1.1) the first factor (the likelihood) is objective, while the
second (the prior) is subjective, since it contains our prejudices about which regions of
the parameter space are more "natural" or "expectable". It is desirable that the results
of the analysis are as independent as possible of the chosen prior. This happens if the
data are powerful enough to select a very small region of the parameter space, so that
eq. (1.1) is dominated by the likelihood, i.e. essentially the pdf is non-zero just in the
narrow region of non-vanishing $p(\text{data}|p_i)$. However, in many instances this is not the
case, as it happens for the MSSM.

The somewhat subjective character of the prior, $p(p_i^0)$, has often motivated to ignore its
presence, identifying in practice $p(p_i^0|\text{data})$ with $p(\text{data}|p_i^0)$. However, it must be noticed
that this procedure implicitly implies a choice for the prior, namely a completely flat prior
in the parameters. This is not necessarily the most reasonable or "free of prejudices"

---

[4]Normally this determination takes the form of a probability distribution since the theoretical com-
putations and the experimental data are affected by different kinds of errors and uncertainties.

attitude. Note for example that using p_i^2 as initial parameters instead of p_i the previous flat prior becomes non-flat. So one needs some theoretical basis to establish, at least, the parameters whose prior can be reasonably taken as flat.

If we are interested in the most probable value of one (or several) of the initial parameters, say p_i, $i = 1, ..., N_1$, but not in the others, p_i, $i = N_1 + 1, ..., N$, we have to *marginalize* the latter, i.e. integrate in the parameter space:

$$p(p_i, \ i = 1, ..., N_1 | \text{data}) \ = \int dp_{N_1+1}, ..., dp_N \ p(p_i, \ i = 1, ..., N | \text{data}) \ . \qquad (1.2)$$

This procedure is very useful and common to make predictions about the values of particularly interesting parameters. It must be noticed that, in order to perform the marginalisation, we need an input for the prior functions *and* for the range of allowed values of the parameters, which determines the range of the definite integration (1.2). A choice for these ingredients is therefore inescapable in trying to make LHC forecasts.

Let us now particularize these general statements to the MSSM (for a review see [10]). Beside the Standard Model (SM) –like parameters (to be discussed below), the MSSM contains a great number of parameters associated with the unknown process of supersymmetry (SUSY) breaking, the so-called soft SUSY-breaking terms. Assuming universality of these terms at a given high scale (namely the scale at which the SUSY breaking is transmitted to the observable sector), these parameters are reduced to four: the universal scalar mass, m, the universal gaugino mass, M, the universal trilinear scalar coupling, A, and the bilinear scalar coupling, B. The universality assumption is in part justified by the need of keeping the FCNC processes under control and it does come out naturally in several schemes of SUSY breaking mediation, e.g. minimal SUGRA or gauge-mediated models (for a review see [11] and [12] respectively). Beside these four parameters one has to include the μ-parameter (i.e. the Higgs mass term in the superpotential) as an additional independent parameter, presumably with a magnitude similar to the soft breaking terms, as it is demanded by a successful electroweak breaking (see below). The notation used here is consistent with refs. [10, 13].

The SM-like parameters of the MSSM include the $SU(3) \times SU(2) \times U(1)_Y$ gauge couplings, g_3, g, g', and the Yukawa couplings, which in turn determine the fermion masses and mixing angles. An important difference from the SM is that the MSSM contains two Higgs doublets, H_1, H_2, with expectation values $v_i = \langle H_i^0 \rangle$ determined by the parameters of the model upon minimization of the scalar potential, $V(H_1, H_2)$. They have to fulfill $2(v_1^2 + v_2^2) = v^2 = (246 \text{ GeV})^2$. The down-type-quark masses go like $m_d \sim y_d v_1 = y_d v \cos \beta$, where $\tan \beta \equiv v_2/v_1$. Similarly for the up-type-quarks $m_u \sim y_u v_2 = y_u v \sin \beta$, and for the charged leptons, $m_e \sim y_e v_1 = y_e v \cos \beta$. Hence the values of the Yukawa couplings which give the observed fermion masses depend on the derived parameter $\tan \beta$, a fact that will be relevant later in our discussion.

In sect. 2 we address some basic aspects of the Bayesian approach for the MSSM, showing in particular that a penalization of the fine-tuning arises from the Bayesian analysis itself (with no ad hoc assumptions as in previous analyses), upon the marginalization of the $\mu-$parameter (subsect. 2.1). We also present a rigorous treatment of the Yukawa couplings, showing that the usual practice of taking the Yukawas "as required",

approximately corresponds to taking logarithmically flat priors in the Yukawa couplings (subsect. 2.2). In sect. 3 we use an efficient set of variables to scan the MSSM parameter space, trading in particular B by $\tan \beta$, giving the effective prior in the new parameters. Finally, in sect. 4 we summarize our results and conclusions.

2   Some basic aspects

2.1   Connection between the Bayesian approach and the fine-tuning measure

It is common lore that the parameters of the MSSM, $\{m, M, A, B, \mu\}$, should not be far from the electroweak scale in order to avoid unnatural fine-tunings to obtain the correct scale of the electroweak breaking. This can be easily appreciated from the minimization of the tree-level form of the scalar potential, $V(H_1, H_2)$, which gives the expectation values of the Higgses, and thus the value of $M_Z^2 = \frac{1}{2}(g^2 + g'^2)(v_1^2 + v_2^2)$; namely

$$M_Z^2 = 2 \frac{m_{H_1}^2 - m_{H_2}^2 \tan^2 \beta}{\tan^2 \beta - 1} - 2\mu^2 \quad . \tag{2.1}$$

Unless the $\mu-$term and the soft masses m_{H_i} (which upon the renormalization running depend also on the other soft terms) are close to the electroweak scale, a funny cancellation among the various terms in the right hand side of (2.1) is necessary to get the experimental M_Z.

A conventional measure of the degree of fine-tuning is given by the Barbieri-Giudice fine-tuning parameters [14]:

$$c_i = \left| \frac{\partial \ln M_Z^2}{\partial \ln p_i} \right|, \tag{2.2}$$

which weigh up the sensitivity of M_Z with respect to the parameters of the model, p_i. The global measure of the fine-tuning is taken as $c \equiv \max\{c_i\}$ or $c \equiv \sqrt{\sum c_i^2}$ [14–17].

Previous studies have attempted to incorporate this fine-tuning measure to the Bayesian approach through the prior $p(p_i)$. In particular, in refs. [2, 18] a prior $p(p_i) \propto 1/c$ was proposed[5]. In principle this is not unreasonable since $1/c$ approximately indicates the probability of a cancellation among the various terms contributing to M_Z^2 to give a result $\lesssim (M_Z^{\mathrm{exp}})^2$. This can be intuitively seen as follows. Expanding $M_Z^2(p_i)$ around a point in parameter space that gives the desired cancellation, say $\mathcal{P}^0 \equiv \{p_i^0\}$, up to the linear term in the parameters, one finds that only a small neighborhood $\delta \mathcal{P} \sim \mathcal{P}^0/c$ around this point gives a value of M_Z^2 smaller or equal to the experimental value [15]. Hence, if one assumes that $\mathcal{P}$ could reasonably have taken any value of the order of magnitude of $\mathcal{P}^0$, then only for a small fraction $\sim 1/c$ of this region one gets $M_Z^2 \lesssim (M_Z^{\mathrm{exp}})^2$, thus the rough probabilistic meaning of c.

---

[5]Another prior designed to catch the naturalness criterion has been proposed in ref. [4].

However, though reasonable, the above-mentioned proposals for priors are rather arbitrary, as the very measure of the fine-tuning is. On the other hand, since the naturalness arguments are deep down statistical arguments, one might expect that an effective penalization of fine-tunings should arise from the Bayesian analysis itself, with no need of introducing "naturalness priors" *ad hoc*. This is in fact the case, as we are about to see.

Let us consider M_Z as an experimental data, on a similar foot to the rest of physical observables. Then the total likelihood reads

$$p(\text{data}|s, m, M, A, B, \mu) \;=\; N_Z \; e^{-\frac{1}{2}\chi_Z^2} \; \mathcal{L}_{\text{rest}} \;, \tag{2.3}$$

where s represents the SM-like parameters, $\mathcal{L}_{\text{rest}}$ is the likelihood associated to all the physical observables, except M_Z, and

$$\chi_Z^2 = \left(\frac{M_Z - M_Z^{\text{exp}}}{\sigma_Z} \right)^2 \;, \tag{2.4}$$

where $\sigma_Z \ll M_Z^{\text{exp}}$ is the experimental uncertainty in the Z mass; finally $N_Z = 1/\sqrt{2\pi}\sigma_Z$ is a normalization constant. Let us now use this sharp dependence on M_Z to marginalize the pdf in the $\mu-$parameter, performing a change of variable $\mu \to M_Z$:

$$
\begin{aligned}
p(s, m, M, A, B| \text{ data}) \;&=\; \int d\mu \; p(s, m, M, A, B, \mu | \text{data}) \\
&=\; N_Z \int dM_Z \left[\frac{d\mu}{dM_Z} \right] e^{-\chi_Z^2} \; \mathcal{L}_{\text{rest}} \; p(s, m, M, A, B, \mu) \\
&\simeq\; \mathcal{L}_{\text{rest}} \left[\frac{d\mu}{dM_Z} \right]_{\mu_0} p(s, m, M, A, B, \mu_0) \;.
\end{aligned}
\tag{2.5}
$$

where μ_0 is the value of μ that reproduces the experimental value of M_Z for the given values of $\{s, m, M, A, B\}$. In the last line of (2.5) we have approximated $N_Z \; e^{-\frac{1}{2}\chi_Z^2} \simeq \delta(M_Z - M_Z^{\text{exp}})$. Essentially the same result is obtained by performing the $\mu-$integration in the stationary point approximation. Now, comparing (2.5) to the definition of fine-tuning parameters (2.2), we can write

$$p(s, m, M, A, B| \text{ data}) \;=\; 2\, \mathcal{L}_{\text{rest}} \; \frac{\mu_0}{M_Z} \; \frac{1}{c_\mu} \; p(s, m, M, A, B, \mu_0) \;. \tag{2.6}$$

Several comments are in order here. First, the presence of the fine-tuning parameter, $1/c_\mu$, penalizes the regions of the parameter space with large fine-tuning, as desired. Actually eq. (2.6) is very similar to multiply by hand the initial prior in the parameters by a factor $1/c$, as in ref. [2]. The difference is that here the factor $1/c_\mu$ has not been put by hand: it comes out from the marginalization in μ. Moreover the prior $p(s, m, M, A, B, \mu_0)$ is still undefined. If one takes it as flat, then one gets the same as in ref. [2], but with one factor μ in the numerator (still the regions of large fine-tuning are penalized since c_μ goes parametrically as $\sim \mu^2$). If one takes logarithmically flat priors, i.e. $p(\mu) \propto 1/\mu$, then eq. (2.6) would formally coincide with the procedure of multiplying the theoretical prior $p(s, m, M, A, B)$ by a factor $1/c$. This is reasonable: the usual naturalness criteria

4

implicitly assume that for a given value of one parameter, say $\mu = \mu_0$, the prior probability is distributed around μ_0 [15,17] with a width $\sim \mu_0$ [see the brief discussion in the paragraph after eq. (2.2)]. This is equivalent to assume that the value $\mu = \mu_0$ has a prior probability $\propto 1/\mu_0$. Actually this is the reason why, according to usual fine-tuning arguments, large soft parameters are more unlikely than small ones: for the former the region of the parameter space that produces the observed electroweak scale is much narrower than for the latter, *not* in absolute value, but *compared* to the size of the soft parameters in each case. Assuming flat priors there would be no reason to prefer soft parameters of the electroweak size instead of e.g. order M_{GUT}. The fact that even for flat priors we still get a penalty factor μ/c_μ comes from the assumption of a prior flat in μ instead of μ^2, which is the quantity that appears in the cancellation [see e.g. eq. (2.1)].

We find very satisfactory that the usual parameter to quantify the degree of fine-tuning emerges from the Bayesian approach "spontaneously", not upon subjective assumptions, especially taking into account that there has been much discussion in the literature about its significance and suitability, see e.g. refs. [15–19]. Actually, one gets simply c_μ instead c, as defined in eq. (2.2). Of course there is nothing special with the $\mu-$parameter, except the fact that we have chosen to marginalize it using the experimental information about M_Z, which is the usual practice. Had we chosen to marginalize another parameter, say M, we would have got c_M, but of course at the end the results would be the same.

A convenient way to view eq. (2.6) is to imagine that we start with an MSSM parameter space $\{s, m, M, A, B\}$ where μ has been eliminated using the experimental value of M_Z. Then the pdf appears as the likelihood associated to the experimental information (except M_Z^{exp}) times an effective prior

$$p_{\mathrm{eff}}(s, m, M, A, B) \;=\; 2\,\frac{\mu_0}{M_Z}\,\frac{p(\mu_0)}{c_\mu}\,p(s, m, M, A, B) \;, \qquad (2.7)$$

where for simplicity we have assumed that the prior in μ factorizes from the rest. This means that the initial prior gets multiplied by a factor $2\frac{\mu_0}{M_Z}\,\frac{p(\mu_0)}{c_\mu}$ that carries the fine-tuning penalty. In Fig. 1 we have plotted this factor in representative slices of the $\{s, m, M, A, B\}$ parameter space (using the two basic choices $p(\mu) \propto$ const., $p(\mu) \propto 1/\mu$) for some illustrative and physically relevant cases. In all of them large soft parameters get penalized (except partially for focus-point regions [20,21]). There are no ad hoc assumptions for this result, it just comes out from the value of M_Z^{exp} and the marginalization of μ.

For practical calculations it is useful to have an approximate expression for c_μ. From the tree-level condition (2.1) we see that $c_\mu \sim 2\mu^2/M_Z^2$. Nevertheless, using the approximate analytic formulas discussed in sect. 3, it is possible to write a much more refined expression for c_μ, which we postpone to that section.

2.2 Nuisance variables and the role of the Yukawa couplings

It is common in statistical problems that not all the parameters that define the system are of interest. In the problem at hand we are interested in determining the probability regions
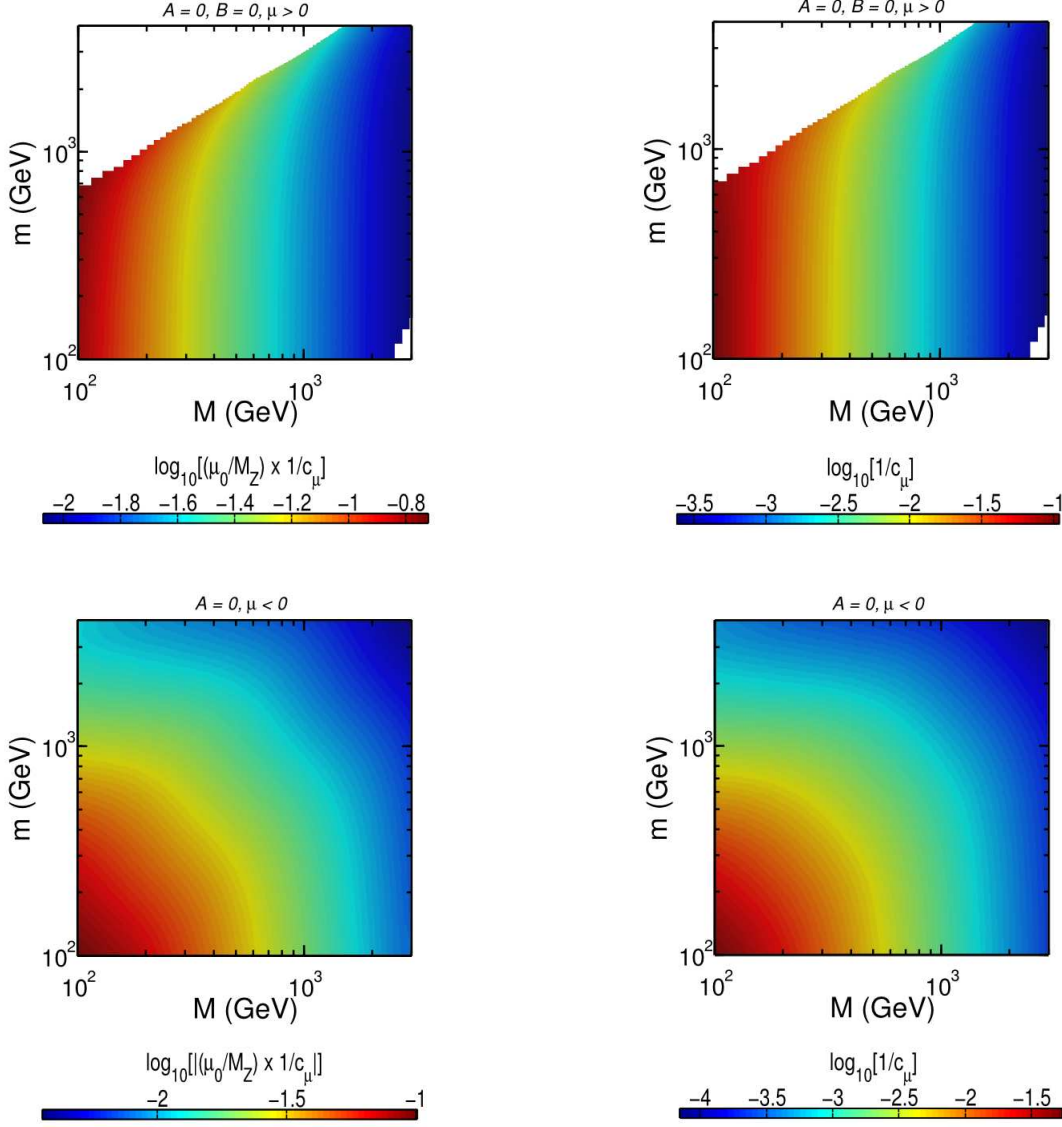
Figure 1: Values of the factor $\mu\, p(\mu)/(M_Z c_\mu)$ (in logarithmic units and up to a convenient proportionality constant) in the $\{m, M\}$ plane for $\mu > 0, A = 0,\ B = 0$ (upper plots), and for $\mu < 0,\ A = 0$ and the minimal SUGRA relation $B = A - m$ (lower plots), using the two basic initial priors, $p(\mu) \propto$ const. (left plots), $p(\mu) \propto 1/\mu$ (right plots). The plotted factor appears in the effective prior given in eq. (2.7).

6

for the MSSM parameters that describe the new physics, i.e. $\{m, M, A, B, \mu\}$, but not (or not at the same level) in the SM-like parameters, denoted by $\{s\}$. However, the *nuisance parameters* $\{s\}$ play an important role in extracting experimental consequences from the MSSM. The usual technique to eliminate nuisance parameters is simply marginalizing them, i.e. integrating the pdf (2.6) in the $\{s\}$ variables (for a review see ref. [22]). When the value of a nuisance parameter is in one-to-one correspondence to a high-quality experimental piece of information (included in $\mathcal{L}_{\mathrm{rest}}$), this integration simply selects the "experimental" value of the nuisance parameter, which thus becomes (basically) a constant with no further statistical significance in the analysis. In particular, the prior on such nuisance parameter becomes irrelevant. In the MSSM, nuisance parameters of this class are the gauge couplings, $\{g_3, g, g'\}^6$, which thus can be extracted from the analysis.

In the pure SM a similar argument can be used to eliminate the Yukawa couplings, since they are in one-to-one correspondence to the quark and lepton masses. However, as discussed in sect. 1, in the MSSM these masses depend also on the value of $\tan \beta \equiv v_2/v_1$, which is a derived quantity that takes different values at different points of the MSSM parameter space. This means that two viable MSSM models (with the same fermion masses) will have in general very different values of the Yukawa couplings, and thus the theoretical prior, $p(y)$, will play a relevant and non-ignorable role in their relative probability. Any Bayesian analysis of the MSSM amounts to an explicit or implicit assumption about the prior in the Yukawa couplings.

In order to make these points more explicit, let us temporarily simplify the discussion approximating the experimental likelihood related to the fermion masses as

$$\mathcal{L}_{\mathrm{fermion\ masses}} \;=\; \delta(m_t - m_t^{\mathrm{exp}}) \, \delta(m_b - m_b^{\mathrm{exp}}) \; .... \tag{2.8}$$

(which is a fair approximation). This is a factor of the global likelihood, $\mathcal{L}_{\mathrm{rest}}$. Likewise, let us approximate the theoretical values of the fermion masses as

$$m_t = \frac{1}{\sqrt{2}} y_t^{\mathrm{low}} v s_\beta, \quad m_b = \frac{1}{\sqrt{2}} y_b^{\mathrm{low}} v c_\beta, \qquad \text{etc.} \tag{2.9}$$

where $s_\beta \equiv \sin \beta$, $c_\beta \equiv \cos \beta$ and y_i^{low} are the low-energy Yukawa couplings. As it is well-known these expressions correspond to the running masses. The physical (pole) masses include a radiative correction that we have ignored here, but not in our full analysis. A further simplification is to assume $y_i^{\mathrm{low}} = R_i y_i$, where y_i are the high-energy Yukawa couplings (and thus the input parameters) and the renormalization-group factor R_i does not depend on y_i itself (this is not a good approximation for the top Yukawa coupling, but we will assume it momentarily for the sake of clarity). Now, the marginalization in the Yukawa couplings can be readily done, integrating the pdf given by eq. (2.6) in the y_i

---

[6]Strictly speaking, the initial theoretical inputs are the gauge couplings at high energy, which are related to the experimental (low-energy) ones by the renormalization-group running. This running depends on the other MSSM parameters through the position of thresholds associated with different particles. Hence, two viable MSSM models have slightly different values of the gauge couplings at high energy, and thus the theoretical prior on the couplings would play an (almost insignificant) role in the statistical comparison of the two models.

variables. Writing just the relevant terms we get

$$\int [dy_t \ dy_b \cdots] \ p(y, m, M, A, B| \ \text{data}) \ = \ \int [dy_t \ dy_b \cdots] \ p(y)\delta(m_t - m_t^{\text{exp}}) \ \delta(m_b - m_b^{\text{exp}})\cdots$$

$$\sim \ p(y) \left|\frac{dy_t}{dm_t}\right| \left|\frac{dy_b}{dm_b}\right| \cdots \ = \ p(y) \ s_\beta^{-1} \ c_\beta^{-1} \cdots \tag{2.10}$$

where $p(y)$ denotes the prior in the Yukawa couplings (which we assume that factorizes from the other priors). Eq. (2.10) represents the footprint of the Yukawa couplings in the pdf. Note that the factors $s_\beta^{-1} \ c_\beta^{-1} \cdots$ arise from the change of variables $y_i \to m_i$, even if the likelihood is not approximated by deltas. There are as many such factors as quarks and leptons. This amounts to a dramatic modulation of the relative probability of MSSM regions with different $\tan\beta$ if one chooses a flat prior, $p(y) = \text{const}$. If, instead, one takes logarithmically flat priors, i.e. $p(y_i) \propto 1/y_i$, then the $s_\beta^{-1} \ c_\beta^{-1} \cdots$ factors get cancelled, so that the elimination of the Yukawa couplings does not leave a footprint in the probability density of the (non-nuisance) MSSM parameter space, $\{m, M, A, B, \mu\}$.

In previous Bayesian analyses of the MSSM the role of the Yukawa couplings was not considered to this extent. Essentially, their values were taken as needed to reproduce the experimental fermion masses, within uncertainties. As we have seen, this practice approximately corresponds to assuming logarithmically flat priors in the Yukawa couplings[7].

The above discussion is however oversimplified. As already mentioned, the marginalization in the top Yukawa coupling (and sometimes the bottom one) produces extra factors due to the dependence of R_t on y_t. Actually, since one is marginalizing simultaneously in the Yukawa couplings and the $\mu-$parameter one has to evaluate the full Jacobian of the transformation $\{\mu, y_t\} \to \{M_Z, m_t\}$, which introduces additional contributions. Furthermore, the picture gets more complicated due to the fact that, for a given choice of $\{m, M, A, B\}$, there may be several values of μ leading to the correct value of M_Z with different values of $\tan\beta$ and thus of the Yukawa couplings. This means that in the marginalization one has to sum over all these possibilities. This is technically annoying and reduces the clarity of the approach. These drawbacks can be eliminated by trading in the statistical analysis the initial $B-$parameter by the derived $\tan\beta$ parameter, as we discuss in the next section.

Let us finally mention that in the analysis of ref. [2] the fermion masses themselves, rather than the Yukawa couplings, were taken as SM-like variables. The advantage of such procedure is that these nuisance variables are in obvious one-to-one correspondence to the experimental data. Then the priors on the masses become almost irrelevant, and they can be integrated out, almost without leaving any footprint. However, this has two problems. First, the fermion masses are obviously derived quantities and should not be taken as initial input variables, even if this makes life easier. Second, such procedure introduces completely artificial factors, as it will become clear at the end of the next section.

---

[7]Actually, for independent reasons, we find the logarithmically flat prior for Yukawa couplings a most sensible choice. Certainly there is no convincing origin for the experimental pattern of fermion masses, and thus of Yukawa couplings. However it is a fact that these come in very assorted orders of magnitude (from $\mathcal{O}(10^{-6})$ for the electron to $\mathcal{O}(1)$ for the top), suggesting that the underlying mechanism may produce Yukawa couplings of different orders with similar efficiency.

3 Efficient variables to scan the MSSM parameter space

In MSSM analyses it is normally very advantageous, both for theoretical and phenomeno-logical reasons, to trade the initial $B-$parameter by the derived $\tan\beta$ parameter. On the phenomenological side, $\tan\beta$ is a parameter that appears explicitly in the predictions for many physical processes, such as cross sections, branching ratios, etc. (this is unlike B, that enters only in a very indirect way). Thus it is convenient to get the probability density of the MSSM parameter space as a function of $\tan\beta$. On the theoretical side, for a given viable choice of $\{m, M, A, \tan\beta\}$, there are exactly two values of μ (with opposite sign and the same absolute value at low energy) leading to the correct value of M_Z. Thus working in one of the two (positive and negative) branches of μ, each point in the $\{m, M, A, \tan\beta\}$ space corresponds exactly to one model, whereas a point in the $\{m, M, A, B\}$ space may correspond to several models, introducing a conceptual and technical complication in the analysis, as mentioned in the previous section.

Changing variables $B \rightarrow \tan\beta$ amounts to a factor $dB/d\tan\beta$ in the pdf. On the other hand, we have seen in sect. 2 that it is convenient to trade μ and y_t by M_Z and m_t, as this makes the marginalization of these variables easier and more transparent. Thus we should compute the whole Jacobian, J, of the transformation

$$\{\mu, y_t, B\} \rightarrow \{M_Z, m_t, t\}, \qquad t \equiv \tan\beta \ , \tag{3.1}$$

so that, in the new variables, the pdf reads

$$p(g_i, m_t, m, M, A, \tan\beta|\ \text{data}) \ = \ \mathcal{L}_{\text{rest}} \ J|_{\mu=\mu_0} \ p(g_i, y_t, m, M, A, B, \mu=\mu_0) \ . \tag{3.2}$$

Here we have made explicit the dependence on the gauge couplings, and the top Yukawa coupling and mass, but not on the other fermions'. In this equation we have already marginalized M_Z using the associated likelihood $\sim \delta(M_Z - M_Z^{\text{exp}})$ (recall that μ_0 is the value of μ that reproduces the experimental M_Z.) The combination

$$p_{\text{eff}}(g_i, m_t, m, M, A, \tan\beta) \ \equiv \ J|_{\mu=\mu_0} \ p(g_i, y_t, m, M, A, B, \mu=\mu_0) \tag{3.3}$$

can be viewed as the effective prior in the new, more convenient, variables to scan the MSSM. Note that, as discussed in subsect. 2.2, the gauge couplings are fairly irrelevant for the statistical analysis, so we will drop them in what follows. In order to work out J we need the dependence of the old variables on the new ones, which can be derived from the minimization equations of the scalar potential, $V(H_1, H_2)$, and from the expression of the top pole mass. For the numerical analysis we have used the SOFTSUSY code [13] which implements the full one-loop contributions and leading two-loop terms to the tadpoles for the electroweak symmetry breaking conditions with parameters running at two-loops. This essentially corresponds to the next-to-leading log approximation. However, in order to highlight the most relevant facts it is useful to write down the expressions arising from the minimization of the tree-level potential with parameters running at one-loop (i.e. essentially the leading log approximation):

$$\mu_{\text{low}}^2 = \frac{m_{H_1}^2 - m_{H_2}^2 t^2}{t^2 - 1} - \frac{M_Z^2}{2} \tag{3.4}$$

$$B_{\text{low}} = \frac{s_{2\beta}}{2\mu_{\text{low}}}(m_{H_1}^2 + m_{H_2}^2 + 2\mu_{\text{low}}^2) \qquad (3.5)$$

$$y_{\text{low}} = \frac{m_t}{v \; s_\beta} \; . \qquad (3.6)$$

Here the "low" subscript indicates that the quantity is evaluated at low scale (more precisely, at a representative supersymmetric mass, such as the geometric average of the stop masses). The soft masses $m_{H_i}^2$ are also understood at low scale. For notational simplicity, we have dropped the subscript t from the Yukawa coupling. We are not making explicit the role of the bottom Yukawa coupling, which is treated in a similar foot to the top one. Note that all these low-energy quantities contain an implicit dependence on the top Yukawa coupling through the corresponding renormalization-group equations (RGEs). The effect of the one-loop corrections on the effective potential to the previous expressions is incorporated by correcting the soft masses $m_{H_i}^2$ with one-loop tadpole effects along the lines of ref. [23]. Similarly the pole top mass is given by the running top mass, appearing in eq. (3.6), plus a radiative correction $\Delta_{rad}m_t$. Eqs. (3.4–3.6), even when corrected with the mentioned radiative effects, have the structure

$$\mu = f(M_Z, y, t), \quad y = g(M_Z, m_t, t), \quad B = h(\mu, y, t) \; , \qquad (3.7)$$

where we only make explicit the dependence on the variables involved in the change of variables (3.1). Note that y depends on M_Z since $v \propto M_Z$. Notice also that, unlike eqs. (3.4–3.6), eqs. (3.7) are defined in terms of the the high-energy parameters.

From eqs. (3.7) it is straightforward to evaluate the Jacobian J of the transformation (3.1), and thus the effective prior (3.3). J gets simply

$$J = \begin{vmatrix} \frac{\partial \mu}{\partial M_Z} & \frac{\partial \mu}{\partial t} & \frac{\partial \mu}{\partial m_t} \\ \frac{\partial B}{\partial M_Z} & \frac{\partial B}{\partial t} & \frac{\partial B}{\partial m_t} \\ \frac{\partial y}{\partial M_Z} & \frac{\partial y}{\partial t} & \frac{\partial y}{\partial m_t} \end{vmatrix} = \frac{\partial f}{\partial M_Z} \frac{\partial g}{\partial m_t} \frac{\partial h}{\partial t} \; , \qquad (3.8)$$

where the factor $\partial f/\partial M_Z$ carries essentially the fine-tuning penalization discussed in subsect. 2.1.

We can give an analytical and quite accurate expression of J by using the approximate equations (3.4–3.6), and expressing the low-energy values of μ, B, y in terms of the high-energy ones through the integrated 1-loop RGEs. Schematically,

$$\mu_{\text{low}} = R_\mu(y)\mu, \quad B_{\text{low}} = B + \Delta_{RG}B(y) \; , \qquad (3.9)$$

where $R_\mu(y), \Delta_{RG}B(y)$ are definite functions of y (and other parameters, but not μ and B) [24]. Similarly,

$$y_{\text{low}} \simeq \frac{yE(Q_{\text{low}})}{1 + 6yF(Q_{\text{low}})} \; , \qquad (3.10)$$

10

where Q is the renormalization scale, $F = \int_{Q_{\text{high}}}^{Q_{\text{low}}} E \ln Q$, and $E(Q)$ is a definite function that depends just on the gauge couplings [25]. Plugging (3.9) and (3.10) into eqs. (3.4–3.6) we get explicit expressions for the f, g, h functions. The relevant derivatives, to be plugged in (3.8), read

$$\frac{\partial f}{\partial M_Z} = -\frac{M_Z}{\mu}\frac{1}{2R_\mu^2} = -\frac{M_Z}{\mu_{\text{low}}}\frac{1}{2R_\mu} \tag{3.11}$$

$$\frac{\partial h}{\partial t} = B_{\text{low}}\frac{1-t^2}{t(1+t^2)} \tag{3.12}$$

$$\frac{\partial g}{\partial m_t} = \frac{E}{v}\frac{1}{s_\beta}\left(\frac{y}{y_{\text{low}}}\right)^2. \tag{3.13}$$

Let us comment briefly on these expressions. As mentioned above, eq. (3.11) is essentially the fine-tuning factor $2\mu/(M_Z c_\mu)$ obtained in subsect. 2.1 [eq. (2.6)]. It penalizes large scales for μ. Eq. (3.12) counts the volume conversion from dB to dt and it is proportional to a soft mass just for dimensional reasons. Note that this factor penalizes low scales. This is easy to understand looking at eq. (3.5): for a given interval in $\tan\beta$, the larger the values of the soft masses and μ, the larger the corresponding interval in B is. So larger B is favoured. Note, however, that the size of the interval of B relative to the value of B itself (which is statistically meaningful) is essentially constant. Indeed, the B-factor in eq. (3.12) will be cancelled in the pdf if one uses logarithmic flat priors for the soft terms, $p(B) \propto 1/B$. This reasoning is similar to that after eq. (2.6). Finally, eq. (3.13) corresponds to eq. (2.10) of our preliminar discussion. In particular, the $1/s_\beta$ factor corresponds to the same factor in (2.10).

The Jacobian of the transformation (3.1) is given by the product of the three factors of eqs. (3.11–3.13),

$$J = \frac{1}{4}(g^2 + g'^2)^{1/2}\left[\frac{E}{R_\mu^2}\right]\frac{B_{\text{low}}}{\mu}\frac{t^2-1}{t(1+t^2)}\left(\frac{y}{y_{\text{low}}}\right)^2 s_\beta^{-1}. \tag{3.14}$$

In the previous derivation we have considered just the top Yukawa coupling in the change of variables (3.1). Once the others fermions are taken into account, the Jacobian gets a s_β^{-1} factor for each $u-$type quark and a c_β^{-1} factor for each $d-$type quark and charged lepton, as discussed in subsect. 2.2. Now, recall that the effective prior in the new variables is the product of J by the initial prior, as expressed in eqs. (3.2, 3.3); so taking a logarithmically flat prior for the Yukawa couplings (i.e. $p(y_i) \propto y_i^{-1}$) the $s_\beta^{-1}, c_\beta^{-1}$ factors get cancelled in the effective prior and the pdf. For the top Yukawa coupling (and sometimes for the bottom one) this cancellation still leaves a residual dependence on $\tan\beta$ since $\left(\frac{y}{y_{\text{low}}}\right)^2 s_\beta^{-1} \times \frac{1}{y} \propto \frac{y}{y_{\text{low}}}$, which through (3.10) depends on y itself and thus on $\tan\beta$.

11

Therefore, the effective prior defined by eq. (3.3) takes the approximate form

$$p_{\text{eff}}(m_t, m, M, A, \tan\beta) \quad \propto \quad \left[\frac{E}{R_\mu^2}\right] \frac{y}{y_{\text{low}}} \frac{t^2 - 1}{t(1 + t^2)} \frac{B_{\text{low}}}{\mu_0} p(m, M, A, B, \mu = \mu_0) \ . \quad (3.15)$$

The most basic priors for the initial variables are the flat and the logarithmic ones, i.e.

$$p(m, M, A, B, \mu) \;=\; \text{const.} \ , \qquad p(m, M, A, B, \mu) \;\sim\; \frac{1}{mMAB\mu} \ . \quad (3.16)$$

Some comments are in order here. First, the normalization factors in (3.16) are determined by the integrated probability and thus depend on the bounds one establishes for the parameters. Since we are discussing here relative probabilities in the parameter space, they are not relevant at this stage, but they become more important when some parameters are marginalized. Second, as argued in subsect. 2.1, the logarithmic prior is physically sensible and is the one that can catch the intuition that fine-tunings are statistically unlikely. Actually, when plugged in (3.15), the logarithmic prior gives rise to the fine-tuning penalization $1/\mu^2 \sim 1/c_\mu$. However, the simple logarithmic prior of eq. (3.15) is clearly too simple, since it cannot be normalized due to low-energy and high-energy divergences. These are easily cured by taking reasonable upper and lower bounds on the parameters, e.g. [10 GeV, M_X]. In fact, this choice can be refined. From the 1-loop RGE of the initial parameters, it is clear that very small values for m, A, B are not radiatively stable, due to sizeable contributions proportional to the gaugino mass M. Therefore, it is not very sensible to assume that values of these parameters smaller than say $\mathcal{O}(10^{-1}M)$ at precisely M_X can have a particular statistical meaning. Thus we can take flat priors at this region of small values. On the other hand, the experimental lower bounds on the gluino, charginos and neutralinos imply that M and μ cannot be smaller than $\mathcal{O}(100)$ GeV.

In Fig. 2 we show the effective prior defined in (3.3) and computed using eq. (3.8) [with the full one-loop expressions of eqs. (3.4–3.6)] for the two priors discussed after eq. (3.16), i.e. flat and logarithmically flat. The plots show, up to a constant of proportionality, the effective priors in the $\{m, M\}$ plane (with constant $\tan\beta, A$) for some representative cases[8]. We have assumed in the figures that the soft terms are initially given at the scale of gauge unification, $M_X \sim 10^{16}$ GeV, as essentially happens in scenarios of gravity–mediated SUSY breaking, but of course our formulas are also applicable to e.g. gauge–mediated SUSY breaking scenarios. The penalization of large scales is clear for the logarithmically flat case, as expected from our discussion. The fact that using a logarithmic prior penalizes large values of the parameters could seem quite obvious. However, this is not so clear when one compares the integrated probability that the the parameters are within different ranges of scales. For instance, the logarithmic prior alone would give more probability to

---

[8]The proportionality constant is simply the normalization constant of the initial prior, eqs. (3.16), times the normalization constant of the Yukawa prior. However, these factors play no role in the exam of the relative probabilities of the points in the parameter space. Obviously, the absolute values of the left plots cannot be compared to those of the right plots, as they are affected by a different normalization constant. We prefer not to include these normalization factors, as they depend on the upper limit assumed for the soft terms, and do not shed any additional light on the relative probabilities inside the parameter space.
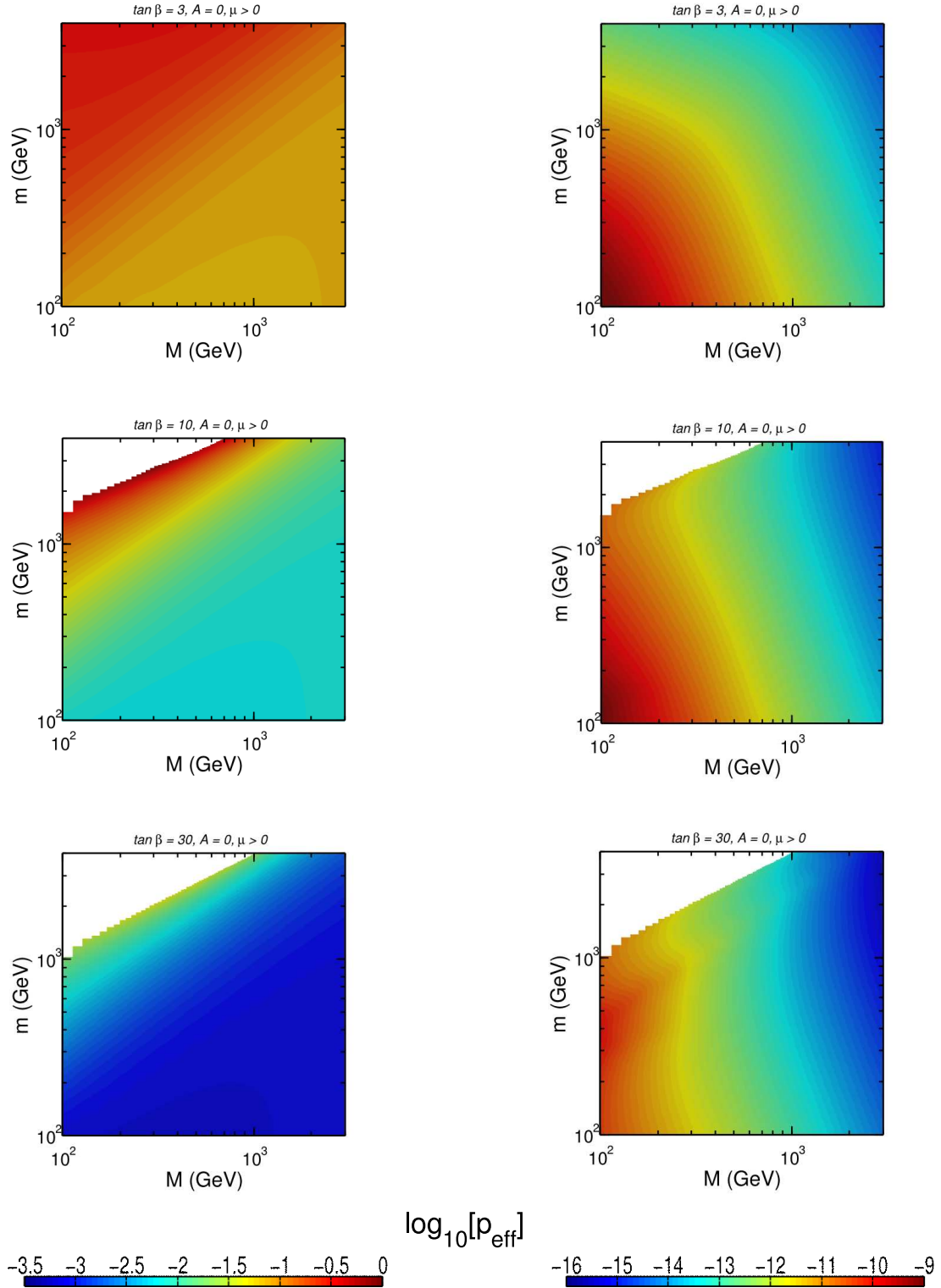
Figure 2: Values of the effective prior, p_{eff}, in logarithmic units as defined in eq. (3.3) (up to a normalization constant), in the $\{m, M\}$ plane for $A = 0$ and $\tan \beta = 3$ (upper plots), $\tan \beta = 10$ (central plots), $\tan \beta = 30$ (lower plots). The left and right plots correspond respectively to the two basic choices of priors (flat and logarithmically flat) discussed in eq. (3.16) and below. See text for further details.
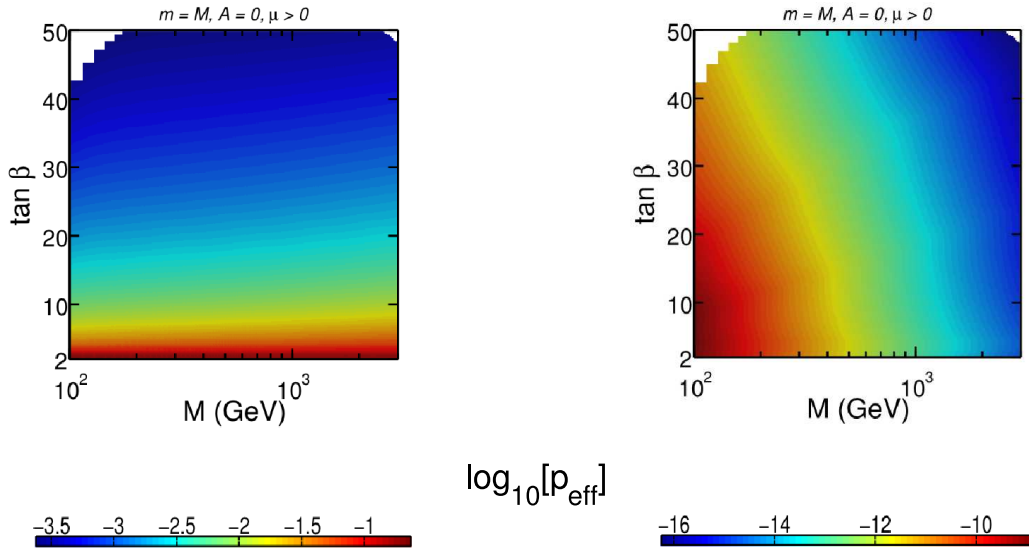
13

Figure 3: The same as Fig. 2, but in the $\{M, \tan\beta\}$ plane, for $A = 0$, $m = M$.

the $[100\text{ TeV}, M_X]$ range than to the $[100\text{ GeV}, 100\text{ TeV}]$ one. However, the presence of the mentioned fine-tuning factor, $1/\mu^2$, in the *effective* prior still penalizes the high-energy regions.

Fig. 3 is similar to Fig. 2, but showing now slices in the $\{M, \tan\beta\}$ plane (with the condition $m = M$). The plots illustrate the $\tan\beta$ dependence of the effective prior, which can be essentially extracted from the approximate expression (3.15). [Note that, besides the explicit dependence, eq. (3.15) contains an implicit dependence on $\tan\beta$ through the R_μ, B_{low} and y/y_{low} factors.] We can appreciate from the plots that the prior probability decreases with $\tan\beta$.

The effective prior computed and shown in the figures corresponds to the last two factors of the pdf (3.2). The first factor, i.e. the likelihood, carries the experimental information (fermion masses, electroweak precision tests, g-2 of the muon, dark matter constraints, etc.). Whatever experimental information (and thus likelihood) we may use, it will be always weighted by the same effective prior factor shown here.

In this section we have argued so far that the sensible initial choice of independent parameters of the MSSM is $\{g_i, y_t, m, M, A, B, \mu\}$, while for practical reasons it is most convenient to work with the set $\{g_i, m_t, m, M, A, \tan\beta, M_Z\}$ (and $\text{sign}\mu$). M_Z is eliminated from the analysis using its extremely sharp likelihood. The effective prior in the new variables is then given by eqs. (3.3, 3.8), for which we gave explicit approximate expressions in eqs. (3.14, 3.15).

It is interesting to wonder what would have been the result if one had insisted in taking directly m_t as an initial (nuisance) variable, so that the transformation (3.1) would have just involved $\{\mu, B\} \rightarrow \{M_Z, t\}$, as has been done e.g. in ref. [4]. As argued in subsect. 2.2, it is theoretically bizarre to take m_t as a fundamental variable, instead of y_t. However, one may gain the bonus of almost no sensitivity to the prior in m_t, since this is

14

essentially fixed by the experiment. This is true, but this procedure introduces extremely counter-intuitive contributions to the Jacobian, as we will see briefly. The new 2-variable Jacobian is given by

$$
J_2 = \begin{vmatrix} \left.\frac{\partial \mu}{\partial M_Z}\right|_{t,m_t} & \left.\frac{\partial \mu}{\partial t}\right|_{M_Z,m_t} \\[2ex] \left.\frac{\partial B}{\partial M_Z}\right|_{t,m_t} & \left.\frac{\partial B}{\partial t}\right|_{M_Z,m_t} \end{vmatrix} , \tag{3.17}
$$

where the subscripts emphasize which variables have to be kept frozen in the partial derivations. Now, using the definitions (3.7), it is straightforward to obtain

$$
J_2 = \frac{\partial f}{\partial M_Z}\frac{\partial h}{\partial t} + \frac{\partial g}{\partial M_Z}\left(\frac{\partial f}{\partial y}\frac{\partial h}{\partial t} - \frac{\partial f}{\partial t}\frac{\partial h}{\partial y}\right) + \frac{\partial f}{\partial M_Z}\frac{\partial h}{\partial y}\frac{\partial g}{\partial t} . \tag{3.18}
$$

It is amusing that this expression is more complicated than in the 3-variable case, eq. (3.8). This comes from the fact that the derivatives in (3.17) contain contributions coming from the dependence of μ and B on y, which is in turn a function of t and M_Z, eq. (3.6). These contributions were cancelled inside the 3-variable Jacobian thanks to the third row in the matrix of eq. (3.8), but they are not cancelled here and give rise to the second and third terms in eq. (3.18). Note that the first term in (3.18) is similar to the 3-variable Jacobian given by eq. (3.8), whose physical significance (including the information about fine-tuning) was discussed after eq. (3.13). This term goes parametrically as B/μ and was the only one quoted in ref. [4], thus the resemblance of their result to our approximate expression (3.14), except for the RG and s_β^{-1} factors. However the second term goes parametrically as $Bm^2/\mu M_Z^2$, and thus is much more important for large soft terms, which then become strongly favoured (contrary to the intuitive expectatives). Therefore there is no reason to have ignored such term. In consequence, the expressions used in ref. [4] are much closer to using y_t as a fundamental variable with logarithmically flat prior than to using m_t.

Let us finish this section by using the approximate expressions discussed above to give, as advanced at the end of subsect. 2.1, an approximate expression for the fine-tuning parameter c_μ. Recall that this parameter was defined as

$$
c_\mu = \left|\frac{\partial \ln M_Z^2}{\partial \ln \mu}\right|_{y,B} , \tag{3.19}
$$

where the subscript indicates that the partial derivative must be performed at y, B constant. Using eqs. (3.7), c_μ can be written as

$$
c_\mu = \frac{2\mu}{M_Z}\left(\frac{\partial f}{\partial M_Z}\right)^{-1}\left[1 + \frac{\partial f}{\partial t}\frac{\partial h}{\partial \mu}\left(\frac{\partial h}{\partial t}\right)^{-1}\right] , \tag{3.20}
$$

where the right hand side has to be understood in absolute value. As above, using (3.9) and (3.10) we obtain explicit approximated expressions for the f, g, h functions. Then eq. (3.20) reads

$$
c_\mu = 4R_\mu^2\frac{\mu^2}{M_Z^2}\left[1 - \frac{t^2(1+t^2)}{(t^2-1)^3}\frac{m_{H_1}^2 - m_{H_2}^2}{B_{\text{low}}\mu_{\text{low}}}\left(\frac{B_{\text{low}}}{\mu_{\text{low}}} - \frac{4t^2}{1+t^2}\right)\right] . \tag{3.21}
$$

15

Note that the combination $m^2_{H_1} - m^2_{H_2}$ can be easily written in terms of B, μ using eqs. (3.4, 3.5).

4   Conclusions

The start of LHC has motivated an effort to determine the relative probability of the different regions of the MSSM parameter space, taking into account the present (theoretical and experimental) wisdom about the model. These attempts are often called "LHC forecasts" [1–4, 6–8]. The central equation to extract this valuable information is the fundamental Bayesian relation

$$p(s, m, M, A, B, \mu | \text{data}) \;\propto\; \mathcal{L}(s, m, M, A, B, \mu) \, p(s, m, M, A, B, \mu) \;, \tag{4.1}$$

which gives this probability in terms of the usual experimental likelihood, $\mathcal{L}$, and the prior $p(s, m, M, A, B, \mu)$, i.e. the "theoretical" probability density assigned a priory to points in the space spanned by the MSSM parameters $\{m, M, A, B, \mu\}$ and the SM-like ones (s).

Since the present experimental data are not powerful enough to select a small region of the MSSM parameter space, the choice of a judicious prior becomes most relevant. Indeed, ignoring this amounts to an implicit choice for the prior (which is not always sensible). On the other hand, it is common lore that the parameters of the MSSM, $\{m, M, A, B, \mu\}$, should not be far from the electroweak scale in order to avoid unnatural fine-tunings to obtain the correct scale of the electroweak breaking. Previous studies have attempted to incorporate this reasonable intuition to the Bayesian approach, by choosing a prior that counted (more or less explicitly) a conventional measure of the fine-tuning, typically the Barbieri-Giudice parameter, c, defined in eq. (2.2).

However, though reasonable, these kinds of proposals are rather arbitrary, as the very measure of the fine-tuning is. On the other hand, since the naturalness arguments are deep down statistical arguments, one might expect that an effective penalization of fine-tunings should arise from the Bayesian analysis itself. One of the main results of this paper has been to show that this is really so: using the fact that the likelihood associated to the experimental M_Z is essentially a Dirac delta, $\sim \delta(M_Z - M_Z^{\text{exp}})$, one can easily marginalize the μ-parameter (i.e. integrate the density of probability in this variable). Then one gets an effective prior for the remaining parameters

$$p_{\text{eff}}(s, m, M, A, B) \;=\; 2 \, \frac{\mu_0}{M_Z} \, \frac{1}{c_\mu} \, p(s, m, M, A, B, \mu_0) \;, \tag{4.2}$$

which exhibits the fine-tuning penalization. (μ_0 is the value of μ that reproduces the experimental M_Z for the given values of $\{s, m, M, A, B\}$.) Of course this effective prior has to be combined with the experimental likelihood, except the part associated to the Z mass. The initial prior, $p(s, m, M, A, B, \mu)$, can be taken as flat or (preferably) logarithmically flat, as usual. We find very satisfactory that precisely the usual parameter to quantify the degree of fine-tuning emerges in the Bayesian approach "spontaneously", not upon subjective assumptions, especially taking into account that there has been much discussion

in the literature about its significance and suitability. We have completed this analysis by giving an explicit and quite accurate expression for c_μ, see eq. (3.21).

Our second result concerns the treatment of the Yukawa couplings. In previous Bayesian analyses the Yukawas were essentially taken as needed to reproduce the experimental fermion masses, within uncertainties. However, unlike the pure SM, in the MSSM the Yukawa couplings are not in one-to-one correspondence to the quark and lepton masses: they depend also on the value of $\tan\beta$, which is a derived quantity that takes different values at different points of the MSSM parameter space. This means that two viable MSSM models (with the same fermion masses) will have in general very different values of the Yukawa couplings, and thus the theoretical prior, $p(y)$, will play a relevant and non-ignorable role in evaluating their relative probability. Any Bayesian analysis of the MSSM amounts to an explicit or implicit assumption about the prior in the Yukawa couplings. We have made explicit the dependence of the results on such prior and shown that the easiest and usual practice of taking the Yukawas "as required", approximately corresponds to taking logarithmically flat priors in the Yukawa couplings, which on the other hand is not an unreasonable choice at all.

Finally we have repeated this analysis, using a more efficient set of variables to scan the MSSM parameter space. Besides trading μ by M_Z and the Yukawa couplings (in particular the top one) by the fermion masses, it is known that trading B by $\tan\beta$ is highly advantageous. Following similar steps one can arrive to an effective prior in the new parameters:

$$p_{\text{eff}}(g_i, m_t, m, M, A, \tan\beta) \equiv J|_{\mu=\mu_0}\, p(g_i, y_t, m, M, A, B, \mu = \mu_0)\,, \qquad (4.3)$$

where J is the Jacobian of the transformation

$$\{\mu, y_t, B\} \rightarrow \{M_Z, m_t, t\}, \qquad t \equiv \tan\beta \qquad (4.4)$$

(M_Z does not appear in the right hand side of (4.3) since it is marginalized as explained above.) Note that still the initial choice of independent parameters is $\{y_t, m, M, A, B, \mu\}$ (on which the initial priors are defined). It is the change of variables plus the marginalization of M_Z what leads to the above effective prior. We have calculated J both numerically and analytically (in an approximate but quite accurate fashion). The relevant formulas are eqs. (3.8) and(3.14). The last expression is very handful and leads to the effective prior given in eq. (3.15). Whatever experimental information (and thus likelihood) one may use, it will be always weighted by the same effective prior factor calculated (and shown in plots for illustrative cases) here.

We have also discussed the results in comparison with other approaches in the literature, arguing that the present one is conceptually more satisfactory.

Acknowledgements

References

[1] B. C. Allanach and C. G. Lester, Phys. Rev. D **73** (2006) 015013 [arXiv:hep-ph/0507283].

[2] B. C. Allanach, Phys. Lett. B **635** (2006) 123 [arXiv:hep-ph/0601089].

[3] R. R. de Austri, R. Trotta and L. Roszkowski, JHEP **0605** (2006) 002 [arXiv:hep-ph/0602028].

[4] B. C. Allanach, K. Cranmer, C. G. Lester and A. M. Weber, JHEP **0708** (2007) 023 [arXiv:0705.0487 [hep-ph]].

[5] L. Roszkowski, R. Ruiz de Austri and R. Trotta, JHEP **0707**, 075 (2007) [arXiv:0705.2012 [hep-ph]].

[6] O. Buchmueller *et al.*, JHEP **0809** (2008) 117 [arXiv:0808.4128 [hep-ph]].

[7] R. Trotta, F. Feroz, M. P. Hobson, L. Roszkowski and R. Ruiz de Austri, arXiv:0809.3792 [hep-ph].

[8] J. Ellis, arXiv:0810.1178 [hep-ph].

[9] R. Trotta, Contemp. Phys. **49** (2008) 71 [arXiv:0803.4089 [astro-ph]].

[10] S. P. Martin, arXiv:hep-ph/9709356.

[11] H. P. Nilles, Phys. Rept. **110** (1984) 1.

[12] C. F. Kolda, Nucl. Phys. Proc. Suppl. **62** (1998) 266 [arXiv:hep-ph/9707450].

[13] B. C. Allanach, *SOFTSUSY: a C++ program for calculating supersymmetric spectra*, Comput. Phys. Commun. **143** (2002) 305 [arXiv:hep-ph/0104145].

[14] R. Barbieri and G. F. Giudice, Nucl. Phys. B **306** (1988) 63.

[15] P. Ciafaloni and A. Strumia, Nucl. Phys. B **494** (1997) 41 [arXiv:hep-ph/9611204].

[16] J. A. Casas, J. R. Espinosa and I. Hidalgo, JHEP **0411** (2004) 057 [arXiv:hep-ph/0410298].

[17] J. A. Casas, J. R. Espinosa and I. Hidalgo, JHEP **0503** (2005) 038 [arXiv:hep-ph/0502066].

[18] L. Giusti, A. Romanino and A. Strumia, Nucl. Phys. B **550** (1999) 3 [arXiv:hep-ph/9811386].

[19] B. de Carlos and J. A. Casas, Phys. Lett. B **309** (1993) 320 [arXiv:hep-ph/9303291]; G. W. Anderson and D. J. Castano, Phys. Lett. B **347** (1995) 300 [arXiv:hep-ph/9409419]; P. Athron and D. J. Miller, Phys. Rev. D **76** (2007) 075010 [arXiv:0705.2241 [hep-ph]].

[20] J. L. Feng, K. T. Matchev and T. Moroi, Phys. Rev. Lett. **84** (2000) 2322 [arXiv:hep-ph/9908309].

[21] J. L. Feng, K. T. Matchev and T. Moroi, Phys. Rev. D **61** (2000) 075005 [arXiv:hep-ph/9909334].

[22] J. O. Berger, B. Liseo and R. L. Wolpert, "Integrated likelihood methods for eliminating nuisance variables" Statistical Science 1999, Vol. 14, No. 1, 1-28

[23] D. M. Pierce, J. A. Bagger, K. T. Matchev and R. j. Zhang, Nucl. Phys. B **491** (1997) 3 [arXiv:hep-ph/9606211].

[24] See e.g. ref. [10] above.

[25] L. E. Ibanez and C. Lopez, Nucl. Phys. B **233** (1984) 511.